MEDICAL RESEARCH LABORATORY

U. S. Naval Submarine Base
New London

Report on

INVESTIGATION OF THE COMBAT INFORMATION CENTER TEST (NavPers 16680) AS A SELECTION DEVICE FOR ENLISTED SUBMARINE RADAR SCHOOL

BuMed project X-496 (Sub. No. 103)
First Final Report

by
Lt. (jg) N.R. Bartlerr, H(S), USNR

22 Oct 1945

Bureau of Medicine and Surgery
Research Project No. X-496 (Sub. No. 103)

APPROVED: Captain C.W. Shilling, (MC), U.S. Navy, MO-in-C

Report on

INVESTIGATION OF THE COMBAT INFORMATION CENTER TEST (NavPers 16680) AS A SELECTION DEVICE FOR ENLISTED SUBMARINE RADAR SCHOOL.

By

Lt. (jg) N. R. Bartlett, H(S), USNR

Final Report

Bureau of Medicine and Surgery

Research Project No. X-496 (Sub. No. 103)

Determination of norms and validity coefficients on a school grade criterion for the CIC Aptitude Test (NavPers 16980) for submarine enlistedmen.

Medical Research Department
U. S. Submarine Base
New London, Connecticut

22 October 1945

TABLE OF CONTENTS

I. SUMMARY:

This report is an exploration of the suitability of the Combat Information Center Test for selecting enlisted Submarine School graduates for training as radar operators. A method of administration, and score standards were developed. The validity of the test in terms of the training criteria available was then assessed, and the results compared with those for tests of the Navy Basic Battery. The cross-relationships between the latter tests and the sections of the CIC test were also calculated.

General findings are discussed, and it is concluded that the evidence does not warrant adoption of the CIC test for assigning Submarine School graduates for radar training. It is suggested, however, that the uses of the Navy Basic Battery can be extended, if desired, to improve the training performance of the population assigned. The Mechanical Aptitude Test is proposed as a screening test for situations where a large surplus of submarine men over the number that can be trained obtains.

The same conclusion is not applied to other populations. It is felt that data reported by NDRC on the validity of related tests are sufficiently conclusive that the CIC testing may be regarded a practical brief screening device for selecting radar operators for training. The tests of the Navy Basic Battery should also be useful in any large heterogeneous group; but for instances when Basic test scores are not available the CIC test is particularly recommended.

II. STATEMENT OF THE PROBLEM:

This study is an investigation of the suitability of the Combat Information Center Aptitude Test for selecting submarine enlisted men for training as Radar Operators. This particular instrument was designed to select officers for Combat Information Center duty; it was

devised by the Bureau of Naval Personnel on the basis of studies by NDRC.[1] The NDRC studies were directed to the problem of selection for proficiency in radar operation. Since the submarine radar watch involves observation and interpretation of the type for which this test selects officers, the same test suggests itself as a selection device for training. However, the level of test difficulty and the absence of specific submarine validation call for caution. Accordingly, a project was undertaken to investigate the suitability of the scale for submarine enlisted men.

One of the factors determining the suitability of a test is validity. For assessing validity, measures of underway performance in standing a submarine radar watch should be developed. None were developed. However, the watch demands almost the same traits and habits of observation and reporting as in the NDRC studies, and in the absence of objective measures of performance it would be best to accept NDRC data on the validity of similar tests for predicting performance on similar radar watches. The alternative is to accept the best criterion available for its limited value. That alternative was pursued in this study; training grades in short courses of instruction are utilized. Therefore, this paper must be regarded as a limited exploration for supplementing work already done. Indeed, it is because suitability of a test depends upon other factors as well as validity that this project was initiated.

III. PROCEDURE:

Following a preliminary investigation in which time limits and methods of administration were established, the test was administered to all candidates for submarine school, as a part of the routine classification program. Every man was examined before starting school.

---

[1] Various reports by NDRC Project SC-70, NS-146, Applied Psychology Panel: "Selection and Training of Oscilloscope Operators."

After the training courses were completed the scores on each of the parts of the test were correlated with school records. The records investigated were (1) grades in radar school and (2) grades in Basic Submarine School. The radar training grades were of two types--the two components of the final grade. The first component is an academic ("examination") mark and second a laboratory ("practical") figure.

In order to determine whether the test improves the selection that can be achieved with the BuPers test battery, similar coefficients for the various scales of the latter battery were also calculated for the same men. And the relationship of each of the parts of the CIC test to tests of the Battery was computed.

IV. RESULTS AND DISCUSSION:

1. Preliminary Study.

Time limits were investigated and the method for presenting the three parts of the test was established in a first preliminary run involving several examination groups, each of 30 men. It developed that the printed instructions on the test booklet were not adequate for many of the lower grade enlisted men. In particular, the instructions for Test I (Polar Grid Coordinate) and for Test II (Scale Reading) were extremely detailed but somewhat confusing to those with limited capacity for grasping instructions. A more effective method proved to be a class demonstration of simple problems with the answer sheet by the examiner. Such demonstration cuts administration time considerably and better insures that every man understands his task, and accordingly it was adopted as standard procedure.

The BuPers time allowances for officers tended to group the larger number of enlisted men at the lower end of the score scale; and the limits were so brief that very few reached the end. Thus the tests involved the element of speed to a great extent. Furthermore,

such brief intervals permit one to identify the few who obtain a high score in a short time but do not allow much distinction among the large group at the bottom of the scale. Neither the speed nor the distribution characteristics seemed desirable for submarine radar selection. Consideration of factors involved in the radar watch did not suggest a strong emphasis on the speed factor. Rapidity in reading the oscilloscope is necessary but speed is probably not a critical factor in the interpretation required of the enlisted men standing the radar watch. If any choices must be made, then, speed might be minimized in the Relative Movement sections of the test, but be retained in the Polar Grid and Scale Reading sections. And in the second place, inasmuch as the proportion trained to stand radar watches is relatively large, it was considered that the test should screen the bottom of the population. In other words, one needed a screen to weed out the few men making up the least promising part of the population rather than a device to identify the very few men who were most promising.

Time limits were arbitrarily set after study of the answer sheets of 120 men who placed marks where they were working as time intervals were called. It is highly probable that the decisions on time limits, through the resulting type of score distribution and the reduced emphasis on the speed factor, influenced the magnitude of the intercorrelation coefficients. Different results might be obtained with other conditions of administration.

Score means and standard deviations for a subsequent population of 300 are shown in Table I. It should be observed that the men on which these norms are based are somewhat superior to the Navy average on aptitude tests; the mean for the group on each of the tests of the BuPers Basic Battery was approximately one standard deviation above the general Navy average.

TABLE I

Standard data for Submarine enlisted men on the Combat Information Center Test.

Test	Total Possible Score	BuPers Time Limits	Submarine Time Limits	Submarine Mean	Submarine Standard Deviation
1. Polar Grid	45	13	18	18.44	6.50
2. Scale Reading	60	12	15	28.28	12.10
3. Relative Movement	45	25	45	24.76	7.36
4. Sum of P.G. and S.R.	105	25	33	46.77	16.65

2. The value of CIC scores for predicting Radar School grades.

On the basis of rates, and other factors of availability, graduates of Basic Submarine School are assigned to Radar School for two weeks of intensive training in how to stand a submarine radar watch. In general, the students thus available for statistical study are highly selected. The norms cited above are based on candidates for the Basic school and, as already stated, those candidates constitute a select group. In addition, the process of selection and training for Basic School tends to restrict the variability of scores even further.

The training criteria must be accepted with all the limitations to be expected in a brief training course. Two kinds of radar grades are assigned--one a laboratory grade and the other an examination grade. The internal consistency of each, and the intercorrelation of one with the other obviously depend on many administrative factors. However, their statistical value for the purposes of this study can be assessed from the following table. In this table, consistency refers to the reliability of the grades. It is represented by the product-moment coefficient between grades for the first week and grades for the second, corrected by the Spearman-Brown formula to represent the reliability of the mean of the two. The relationship between examination and laboratory marks is the product-moment coefficient of the means.

TABLE II

Consistencies and Interrelationships of two grades in Radar School.

	Class 9,10 and 11	Class 12 and 13.
Examination Grade Consistency	0.56	0.59
Laboratory Grade Consistency	0.59	0.46
Relationship between Examination and Laboratory Grades	0.24	0.26
Number in population	86	63

The relationships between scores on the CIC test and the above laboratory and examination grades are shown in Table III. Classes 9 through 13 were utilized in computing these data.

TABLE III

Product-moment correlation coefficients of CIC scores with grades in Submarine Radar School.

Test Section	Average Examination	Average Laboratory
Part I - Polar Grid	-0.04	0.44
Part II - Scale Reading	0.29	0.24
Part III - Relative Movement	0.36	0.23

In terms of the criteria available in Radar training, then, it appears that the CIC test has some validity. It is interesting that Part III has the most in common with the examinations grade, and Part I with the laboratory grade. Coefficients of this magnitude indicate improvement in the population if minimum scores were used as a basis for assignment; but in view of the doubtful criteria, the evidence above on predictability does not warrant acceptance of the test as a single basis for assignment.

3. The relationship between CIC scores and grades in Submarine School.

A sample of 158 men in one single graduating class was studied in order to determine the relationship between CIC scores and grades in Basic Submarine School. This training measure has been discussed in considerable detail in the report of another project.[2] Coefficients are presented in Table IV.

TABLE IV

Product-moment coefficients of the CIC scores with grades in Basic Submarine School.

Test Section	Coefficient
Part I - Polar Grid	0.21
Part II - Scale Reading	0.23
Part III - Relative Movement	0.29
Sum of Parts I and II	0.25
Sum of Parts I, II and III	0.31

Although the coefficients in Table IV indicate a positive relationship to this criterion of general submarine training, the significant point is that the coefficients are low, and probably less than coefficients with scores for tests of the BuPers Basic Battery. For the same 158 men, the GCT, Arithmetic and Mechanical Knowledge (Electrical) tests each yielded a coefficient exceeding 0.30.

4. Relationship between CIC scores and tests of the BuPers Basic Battery and intercorrelations of CIC scores.

Even if this new test procedure proves to have a significant capacity for forecasting either training performance or operational performance, the question which must be answered is whether that capacity should be utilized. The answer involves not only the problem of how much predictive capacity the test has, but also whether the test has more than other methods and whether the test can be fitted into selection and training time-schedules. In other words, if the test is valid, it may be useful; but whether it should be used depends upon whether it constitutes a worthwhile improvement over the techniques already in use or available for use.

Every candidate for submarine training is required to undergo tests of the BuPers Basic Battery, when scores for those tests are not in his service record. Thus a complete set of battery scores is available for every candidate for school. Since this is a standard procedure that is also in effect at several other naval activities it is most logical to use scores of the BuPers battery for selection whenever that battery is as valid as other methods available which require special testing. Other things being equal, the BuPers battery method must be considered preferable to any requiring special examinations.

The extent to which the various sections overlap with scales of the BuPers battery is shown in Table V below. The forms of the

battery represented are 2 and 3 of the edition designed for Training Centers. In the same table are presented the product-moment coefficients of relationship between the three sections of the test. The number of men involved in these computations was 225.

TABLE V

Correlation Coefficients of Interrelationships among Tests.

		CIC Section	
	Polar Grid	Scale Reading	Relative Movement
Scale Reading	0.55		0.51
Relative Movement	0.38	0.51	
GCT	0.25	0.43	0.50
Reading	0.17	0.30	0.46
Arithmetic	0.28	0.48	0.53
MAT	0.37	0.43	0.49
MK(Mech)	0.17	0.08	0.10
MK(Elec)	0.27	0.41	0.29
Sum of GCT and Arithmetic	0.30	0.52	0.64
Sum of GCT, Arithmetic and MAT	0.38	0.56	0.64

It is obvious from Table V that there is a considerable degree of overlapping between tests of the BuPers battery and the sections of the CIC test. In view of the restricted variability of the sample on which these data are based this finding must be particularly emphasized.

5. The value of BuPers battery scores for predicting Radar School grades.

The relationships between scores of the regular battery of tests inserted in the Service Record and grades in Radar school are illustrated in Table VI for the same population utilized for Table III. Coefficients for the Reading and MK(Mech) tests were not computed; inspection of scattergrams showed these tests to bear little relationship to either grade.

It is noted that the coefficients are of the same general magnitude as those in Table III; for predicting Radar training grades, the BuPers battery is at least as efficient as the CIC test, if not more efficient.

TABLE VI

Product-moment coefficients of BuPers Battery scores with grades in Submarine Radar School.

Test	Average Examination	Average Laboratory
GCT	0.33	0.07
Arithmetic	0.33	0.20
MAT	0.31	0.44
MK(Elec)	0.37	0.35
Sum of MKE and Arithmetic	0.40	0.34
Sum of MKE and MAT	0.38	0.46

V. COMMENTS ON SUITABILITY OF CIC TEST FOR RADAR TRAINING.

1. Selection of submarine school graduates for radar training.

In the light of training grades in both Basic and Radar Submarine schools, the tests of the BuPers battery are probably slightly more efficient than the sections of the CIC test. It is conceded that the training grades investigated are not excellent criteria for training performance, and furthermore that, in any case, indices of training are not measures of adequacy in underway performance. But there is one significant negative finding: at least there is no evidence in this study that the CIC test should supplant tests of the present battery. Then, second, there is one positive finding: even if later evidence showing a high degree of validity for predicting operator efficiency on submarine patrols should develop for the CIC test, the cross correlations with tests of the present battery are great enough to anticipate that a considerable validity for the latter would appear also. Accordingly, it is recommended that the uses of the present battery be extended to cover any selection from Submarine school graduates for Radar training, rather than that a new testing method be established.

GCT scores are one basis for assignment to submarine training, and GCT, Arithmetic and MK(Elec) scores correlate significantly with academic grades in Basic school.[2] Therefore the selection process that will take place before assignment to Radar training will involve the factors in those tests. On the other hand, the MAT test is significantly related to Radar grades, and at the same time appears less involved than

---

1,2 Op.cit.

the other measures in the natural process of elimination before assignment. Furthermore the MAT test is associated to a considerable extent with each of the three parts of the CIC test. Assuming on the basis of related NDRC results[1] that the CIC test is a valid selection device, the MAT probably will also carry at least some validity. Following this argument, then, the MAT test is proposed as a single test for eliminating a part of the Submarine School graduates for Radar training.

No particular cutting scores are recommended. Administrative policies govern both the numbers and rates of men assigned to Radar school, and the numbers and rates assigned to Radar watches aboard. For example, the current policy is to assign no man for Radar training who is not also qualified for Sonar training. It is not the province of this activity to comment on the advisability of such policies; but a selection activity must modify its programs to the personnel situation. It has come to the attention of this activity that all too frequently men are disqualified for training for various watches who later must be used on those watches anyway even though not trained ashore previously. In such a situation rejection limits should be set in conformity with the potential number to be used on the watches, rather than set in conformity with the existing training quotas. Training quotas change with administrative policies; entries in Service and Health Records are not erased so easily. It is noted that a minimum assignment score of 50 rejects approximately five percent of those now asigned for Radar training. If a cut must be suggested, this one is proposed as being of the right

---

[1] Op. cit.

magnitude for the current personnel situation. However, it is emphasized again that whatever cutting scores are adopted must be chosen in the light of administrative and availability factors. This comment is a general one applying to any kind of selection for which validity is limited.

The evidence in this paper on the validity of the CIC test is inconclusive. It is urged, however, that this fact be ignored in the light of the significant validation reported by NDRC. It appears that the CIC test has much in common with the BuPers battery, so to a large extent the two are interchangeable. Perhaps the CIC scale may be especially useful in situations where additional aptitudes tests are required, but for the purpose of choosing submarine school graduates for radar training it is concluded that the introduction of the CIC test into the routine battery is not warranted.

It is suggested that, for a population already highly selected on the basis of paper-and-pencil aptitudes tests, it is in order to search elsewhere than in written tests for ways to improve selection further. For example, aptitudes tests have been utilized to a great extent in the selection of submarine school candidates; but measures of visual functions have not been stressed. When this study was initiated, a measure of acuity for well-illuminated objects seen at a distance was the only measure applied. Rather than to explore the need for further emphasis in the area of written tests it would seem appropriate to investigate such questions as sensory capacity for the task or motor capacity for sustained observation.

2. General Selection for radar training.

The comments in the previous section are limited to one fairly homogeneous population - submarine school graduates. For other populations (including submarine men who have not been selected and trained in the New London Submarine School) the CIC test or some other comprehensive set of tests can be used to eliminate very poor student material before training. Correlations for the homogeneous sample reported in this study are low, but nevertheless they indicate that at least some significant training factors are involved in the CIC test. This finding would have been predicted, of course, from the NDRC work At any rate, the data indicate that if a man makes an extremely poor showing on the aptitude tests there would be a definite question of his potentialities for radar training.

The CIC test is simple to administer, easy to score, and, as used in this study, the total time necessary for testing a small class is about 95 minutes. Furthermore, the scales do not involve a repetition of the Basic Battery, and so probably neither affect test scores appreciably on any subsequent Basic Battery test nor are themselves influenced by previous Basic Battery tests. Thus perhaps the CIC test should be recommended especially for situations where Basic Battery scores are not available.

VI. GENERAL CONCLUSIONS:

1. The Combat Information Center test overlaps considerably with tests of the Navy Basic Battery.

2. The Combat Information Center test is not recommended as a device for further selection of Submarine School graduates for radar training.

3. However, it is felt that the test should be useful for selecting from general Navy populations for radar training.

4. Test methods and statistical data on the CIC test developed in this investigation should apply also to other populations whose average scores on the Navy Basic Battery tests are approximately one standard deviation above the general Navy average.

5. It is proposed that the Mechanical Aptitude test of the Basic Battery be utilized if personnel availabilities indicate the desirability of further selection of submarine radar operators on the basis of written aptitudes tests.